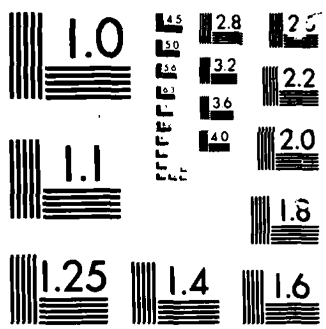
AD-A167 455 A CONVERGENCE THEOREM FOR NEWTON-LIKE METHODS IN BANACH SPACES(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER T YAMAMOTO JAN 86 MRC-TSR-2904 DAAG29-80-C-0041 1/1

UNCLASSIFIED F/G 12/1 NL



M'CROCOP' CHART

(2)

AD-A167 455

MRC Technical Summary Report # 2904

A CONVERGENCE THEOREM FOR NEWTON-LIKE METHODS IN BANACH SPACES

Tetsuro Yamamoto

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

January 1986

(Received December 5, 1985)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

A CONVERGENCE THEOREM FOR NEWTON-LIKE METHODS IN BANACH SPACES

Tetsuro Yamamoto*

Technical Summary Report #2904

January 1986

ABSTRACT

A convergence theorem for Newton-like methods in Banach spaces is given, which improves results of Rheinboldt [25], Dennis [2], Miel [13, 14] and Moret [16] and includes as a special case an updated version of the Kantorovich theorem for the Newton method given in previous papers [33-35]. Error bounds obtained in [32] are also improved. This paper unifies the study of finding sharp error bounds for Newton-like methods under Kantorovich type assumptions.

AMS (MOS) Subject Classifications: 65G99, 65J15

Key Words: A convergence theorem, error estimates, Newton-like methods, Kantorovich type assumptions

Work Unit Number 3 - Numerical Analysis and Scientific Computing

*Department of Mathematics, Faculty of Science, Ehime University, Matsuyama 790, Japan.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041 and by the Ministry of Education, Japan.

SIGNIFICANCE AND EXPLANATION

To find sharp error bounds for iterative solution of nonlinear equations in Banach spaces is of basic importance in numerical analysis. This paper gives a convergence theorem for a class of Newton-like methods in Banach spaces, which improves the theorems of Kantorovich [7, 8], Rheinboldt [25], Dennis [2], Miel [13, 14] and Moret [16]. The argument employed in this paper certainly simplifies and unifies the study for finding sharp error bounds for Newton-like methods in Banach spaces.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

A CONVERGENCE THEOREM FOR NEWTON-LIKE METHODS IN BANACH SPACES

Tetsuro Yamamoto*

1. Introduction

Let X and Y be Banach spaces and consider an operator $F : D \subseteq X \to Y$ which is Fréchet differentiable in an open convex set $D_0 \subseteq D$. Many iterative methods for solving the equation

$$F(x) = 0 \tag{1.1}$$

can be written in the form

$$x_{n+1} = x_n - A(x_n)^{-1}F(x_n) \quad , \quad n \geq 0 \quad , \tag{1.2}$$

where $x_0 \in D_0$ is given and $A(x)$ denotes a linear operator which approximates the Fréchet derivative $F'(x)$ of F. Under some assumptions, Rheinboldt [25] established a convergence theorem for (1.2) which includes the Kantorovich theorem for the Newton method ($A(x_n) = F'(x_n)$) as a special case. A further generalization was given by Dennis [2, 3]. Miel [13, 14] improved the error bounds for Rheinboldt [25]. Moret [16] obtained a convergence theorem as well as error bounds for the iteration (1.2) under the stronger conditions than those of Rheinboldt. By numerical examples, he showed that his bounds are sharper than those of Miel. *However, no proof* is given. Recently, in [32], we presented a method for finding sharp error bounds for (1.2) under Dennis' assumptions and showed that the bounds obtained improve those of Rheinboldt, Dennis and Miel and reduce to Moret's bounds if we replace the assumptions by his stronger ones. It was also shown that Moret's results can be derived from Rheinboldt's.

In this paper, we first state an updated version of the Kantorovich theorem for the Newton ethod in §2. Next, in §3, we give a simple but powerful principle for finding error bounds for (1.2) under Kantorovich type assumptions. Finally, as an application of this prin^ciple, a convergence theorem for (1.2) is given in §4, which includes the updated

*Department of Mathematics, Faculty of Science, Ehime University, Matsuyama 790, Japan.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041 and by the Ministry of Education, Japan.

version of the Kantorovich theorem and improves the results of Rheinboldt, Dennis, Miel and Moret. Our approach certainly simplifies and unifies the study for finding sharp error bounds for the Newton-like methods. (Also see Yamamoto [32-35].)

2. An Updated Version of the Kanotorovich Theorem

Let F, D_0 and x_0 be defined as in §1 and $F(x_0) \neq 0$ without loss of generality. Furthermore, assume that $F'(x_0)^{-1}$ exists and the following conditions are satisfied:

$$\|F'(x_0)^{-1}(F'(x) - F'(y))\| \leq K\|x-y\|, \; x,y \in D_0, \; K > 0 \quad ,$$

$$\eta = \|F'(x_0)^{-1}F(x_0)\|, \; h = k\eta \leq \frac{1}{2}, \; t^* = \frac{2\eta}{1+\sqrt{1-2h}} \quad ,$$

$$\bar{S} = \bar{S}(x_1, t^*-\eta) = \{x \in X \mid \|x-x_1\| \leq t^*-\eta\} \subseteq D_0 \quad .$$

Under these assumptions, define the scaler sequence $\{t_n\}$ by

$$t_0 = 0, \; t_{n+1} = t_n - f(t_n)/f'(t_n), \; n \geq 0 \quad ,$$

with $f(t) = \frac{1}{2}Kt^2 - t + \eta$, and the sequences $\{B_n\}$, $\{\eta_n\}$ and $\{h_n\}$ by

$$B_0 = 1, \; \eta_0 = \eta, \; h_0 = h = K\eta \quad ,$$

$$B_n = B_{n-1}/(1 - h_{n-1}), \; \eta_n = \frac{1}{2}KB_n\eta_{n-1}^2, \; h_n = KB_n\,\eta_n, \; n \geq 1 \quad .$$

Then, in [34, 35], we obtained the following result, which is an updated version of the Kantorovich theorem and essentially equivalent to [33, Theorem 3.1] with the optimal parameter φ.

Theorem 2.1. With the above notation and assumptions, we have the following:
(i) The Newton method $x_{n+1} = x_n - F'(x_n)^{-1}F(x_n)$ is well defined for every $n \geq 0$, $x_n \in S$(interior of $\bar{S}$) for $n \geq 1$ and $\{x_n\}$ converges to a solution $x^* \in \bar{S}$ of the equation (1.1).
(ii) The solution x^* is unique in

$$\tilde{S} = \begin{cases} S(x_0, t^{**}) \cap D_0 & (2h < 1) \\ \bar{S}(x_0, t^{**}) \cap D_0 & (2h = 1) \end{cases} \quad ,$$

where $t^{**} = (1 + \sqrt{1-2h})/K$ and $S(x_0, t^{**})$ denotes the interior of $\bar{S}(x_0, t^{**})$.

(iii) Let $\bar{S}_0 = \bar{S}$, $\bar{S}_n = \bar{S}(x_n, t^*-t_n)\,(n \geq 1)$,

$$K_n = \sup_{\substack{x,y\in\bar{S}_n \\ x\neq y}} \frac{\|F'(x_n)^{-1}(F'(x) - F'(y))\|}{\|x - y\|} \quad (n \geq 0) \quad ,$$

$$L_n = \sup_{\substack{x,y\in\bar{S} \\ x\neq y}} \frac{\|F'(x_n)^{-1}(F'(x) - F'(y))\|}{\|x - y\|} \quad (n \geq 0) \quad .$$

Then, for $n \geq 0$, we have $t_{n+1} - t_n = \eta_n$ and the following error estimates hold:

$$\|x^* - x_n\| \leq \frac{2d_n}{1 + \sqrt{1-2K_n d_n}}$$

$$\leq \frac{2d_n}{1 + \sqrt{1-2L_n d_n}} \tag{2.1}$$

$$\leq \frac{2d_n}{1 + \sqrt{1-2K(1-K\Delta_n)^{-1} d_n}} \quad \text{(Moret [16])} \tag{2.2}$$

$$\leq \frac{2d_n}{1 + \sqrt{1-2KB_n d_n}} \quad \text{(Kantorovich [7])}$$

$$\leq \frac{2d_n}{1 + \sqrt{1-2h_n}} \quad \text{(Döring [5])}$$

$$= \frac{t^*-t_n}{\nabla t_{n+1}} d_n \quad \text{(Miel [14, 15])}$$

$$\leq t^* - t_n \quad \text{(Kantorovich [8])}$$

$$= \frac{2\eta_n}{1 + \sqrt{1-2h_n}} \quad \text{(Kantorovich [7])}$$

$$\leq \frac{1}{2^{n-1}} (2h)^{2^n-1} \eta \quad \text{(Kantorovich [7])} \quad ,$$

where $d_n = \|x_{n+1} - x_n\|$, $\Delta_n = \|x_n - x_0\|$ and ∇ denotes the backward difference operator.

The well known bounds obtained by Dennis [1], Rall-Tapia [24], Tapia [28], Ostrowski [18, 19], Gragg-Tapia [6] and Potra-Pták [22] fall into the above chart (cf. Yamamoto [31 - 35] and Potra [21]). Furthermore, it was shown in [33 - 35] that (2.1) improves the bounds of Lancaster [11], Kornstaedt [10] and Potra [21]. It is also easy to see that the bound of Potra [20] follows from (2.2). Therefore, Theorem 2.1 gives a unified derivation of the known error bounds for the Newton method under the Kantorovich assumptions.

3. A Principle for Finding Error Bounds for (1.2)

Before extending Theorem 2.1 to the iteration (1.2) we prove the following result.

Theorem 3.1. Let the equation (1.1) have a solution x^* and consider the iteration (1.2). Let x_n and x_{n+1} be defined for some $n \geqq 0$, and a_n, b_n and c_n be nonnegative numbers such that $a_n > 0$, $b_n \leqq 1$ and

$$\|x^* - x_{n+1}\| \leqq \frac{1}{2} a_n \|x^* - x_n\|^2 + b_n \|x^* - x_n\| + c_n \quad .$$

Furthermore, put $d_n = \|x_{n+1} - x_n\|$. If the polynomial

$$\bar{p}(t) = \frac{1}{2} \bar{a}_n t^2 - (1-\bar{b}_n)t + \bar{c}_n + d_n$$

with $\bar{a}_n \geqq a_n$, $1 \geqq \bar{b}_n \geqq b_n$ and $\bar{c}_n \geqq c_n$ has two positive zeroes $\bar{\sigma}_n^*$, $\bar{\sigma}_n^{**}$ such that $\bar{\sigma}_n^* \leqq \bar{\sigma}_n^{**}$, then the polynomial $p(t) = \frac{1}{2} a_n t^2 - (1-b_n)t + c_n + d_n$ also has two positive zeroes σ_n^*, σ_n^{**} such that $\sigma_n^* \leqq \bar{\sigma}_n^* \leqq \bar{\sigma}_n^{**} \leqq \sigma_n^{**}$. If it is known for any reasons that $\|x^* - x_n\| \leqq \bar{\sigma}_n^*$, then we have an improved error estimate $\|x^* - x_n\| \leqq \sigma_n^*$.

Proof. The first assertion of the theorem easily follows from the fact that $\bar{p}(t) \geqq p(t)$ for $t > 0$. To prove the second assertion, we observe that $p(\|x^* - x_n\|) \geqq 0$ since

$$\|x^* - x_n\| - d_n \leqq \|x^* - x_{n+1}\| \leqq \frac{1}{2} a_n \|x^* - x_n\|^2 + b_n \|x^* - x_n\| + c_n \quad .$$

Therefore we have

$$\|x^* - x_n\| \leqq \sigma_n^* \quad \text{or} \quad \|x^* - x_n\| \geqq \sigma_n^{**} \quad .$$

However, if it is known for any reasons that $\|x^* - x_n\| \leqq \bar{\sigma}_n^*$, then the latter case can be excluded: In fact, we have

$$\sigma_n^* < \bar{\sigma}_n^* \leqq \bar{\sigma}_n^{**} < \sigma_n^{**} \quad \text{if} \quad \bar{a}_n + \bar{b}_n + \bar{c}_n > a_n + b_n + c_n$$

and

$$\sigma_n^* = \bar{\sigma}_n^* \leq \bar{\sigma}_n^{**} = \sigma_n^{**} \quad \text{if} \quad \bar{a}_n + \bar{b}_n + \bar{c}_n = a_n + b_n + c_n \quad .$$

Q.E.D.

The usual convergence proof for (1.2) is done with the use of a majorant sequence $\{t_n\}$ due to Rheinboldt [25] such that $\|x_{n+1} - x_n\| \leq t_{n+1} - t_n$ and $t_n \leq t_{n+1} \leq \cdots \to t^*$ as $n \to \infty$. Therefore, by taking $t^* - t_n$ for $\bar{\sigma}_n^*$ in Theorem 3.1, we can apply the theorem to obtain sharper error bounds. The detailed argument will be given in the next section.

4. A Convergence Theorem for (1.2)

Let F, D_0 and x_0 be defined as in §1 and consider the Newton-like method (1.2). According to Dennis [2] and Schmidt [26], we assume the following:

$$\|A(x_0)^{-1}(F'(x) - F'(y))\| \leq K\|x-y\|, \; x,y \in D_0, \; K > 0 \quad ,$$

$$\|A(x_0)^{-1}(A(x) - A(x_0))\| \leq L\|x-x_0\| + \ell, \; x \in D_0, \; L \geq 0, \; \ell \geq 0 \quad ,$$

$$\|A(x_0)^{-1}(F'(x) - A(x))\| \leq M\|x-x_0\| + m, \; x \in D_0, \; M \geq 0, \; m \geq 0 \quad ,$$

$$\ell + m < 1, \; \sigma = \max(1, \frac{L+M}{K}), \; F(x_0) \neq 0 \quad ,$$

$$\eta = \|A(x_0)^{-1}F(x_0)\|, \; h = \sigma K\eta/(1 - \ell - m)^2 \leq 1/2 \quad ,$$

$$t^* = (1 - \ell - m)(1 - \sqrt{1-2h})/(\sigma K) \quad ,$$

$$\tilde{t}^{**} = (1 - m + \sqrt{(1-m)^2 - 2K\eta})/K \quad ,$$

$$\bar{S} = \bar{S}(x_1, t^*-\eta) \subseteq D_0 \quad .$$

Under these assumptions, define the sequence $\{t_n\}$ by

$$t_0 = 0, \; t_{n+1} = t_n + f(t_n)/g(t_n), \; n \geq 0 \quad ,$$

with $f(t) = \frac{1}{2}\sigma K t^2 - (1 - \ell - m)t + \eta$ and $g(t) = 1 - \ell - Lt$, and the sequences $\{p_n\}$, $\{q_n\}$, $\{s_n\}$, $\{\eta_n\}$ and $\{h_n\}$ by

$$p_0 = 1 - \ell,\ q_0 = 1 - \ell - m,\ B_0 = p_0/q_0^2,\ \eta_0 = \eta,\ h_0 = \sigma K B_0 \eta_0 \quad ,$$

$$p_n = 1 - \ell - L \sum_{j=0}^{n-1} \eta_j,\ q_n = 1 - \ell - m - \sigma K \sum_{j=0}^{n-1} \eta_j,\ B_n = p_n/q_n^2 \quad ,$$

$$\eta_n = \{\tfrac{1}{2}\sigma K \eta_{n-1}^2 + (p_{n-1} - q_{n-1})\eta_{n-1}\}/p_n,\ h_n = \sigma K B_n \eta_n,\ n \geq 1 \quad .$$

Furthermore, put

$$\varphi(t) = 1 - \ell - m - (L+M)t,\ \Delta_n = \|x_n - x_0\|$$

and

$$d_n = \|x_{n+1} - x_n\| \quad .$$

Then we can prove the following result, which is a natural generalization of Theorem 2.1.

Theorem 4.1. With the above notation and assumptions, we have the following:

(i) The iteration (1.2) is well defined for every $n \geq 0$, $x_n \in S$(interior of $\bar{S}$) for $n \geq 1$ and $\{x_n\}$ converges to a solution $x^* \in \bar{S}$ of the equation (1.1).

(ii) The solution x^* is unique in

$$\tilde{S} = \begin{cases} S(x_0, \tilde{t}^{**}) \cap D_0 & (\text{if } 2K\eta < (1-m)^2) \\ \bar{S}(x_0, \tilde{t}^{**}) \cap D_0 & (\text{if } 2K\eta = (1-m)^2) \quad . \end{cases} \tag{4.1}$$

(iii) Let $\bar{S}_0 = \bar{S}$, $\bar{S}_n = \bar{S}(x_n, t^* - t_n)(n \geq 1)$,

$$K_n = \sup_{\substack{x,y\in\bar{S}_n \\ x\neq y}} \frac{\|A(x_n)^{-1}(F'(x) - F'(y))\|}{\|x - y\|} (n \geq 0) \quad ,$$

$$L_n = \sup_{\substack{x,y\in\bar{S} \\ x\neq y}} \frac{\|A(x_n)^{-1}(F'(x) - F'(y))\|}{\|x - y\|} (n \geq 0) \quad .$$

Then we have

$$x^* \in \bar{S}_n \subseteq \bar{S}_{n-1} \subseteq \cdots \subseteq \bar{S}_0 \quad ,$$

$$t_{n+1} - t_n = \eta_n,\ 2h_n \leq 1$$

and

$$\|x^* - x_n\| \leq \alpha_n \equiv \frac{2g(\Delta_n)d_n}{\varphi(\Delta_n) + \sqrt{\varphi(\Delta_n)^2 - 2K_n g(\Delta_n)^2 d_n}} \quad (n \geq 0)$$

$$\leq \beta_n \equiv \frac{2g(\Delta_n)d_n}{\varphi(\Delta_n) + \sqrt{\varphi(\Delta_n)^2 - 2L_n g(\Delta_n)^2 d_n}} \quad (n \geq 0)$$

$$\leq \gamma_n \equiv \frac{2g(\Delta_n)d_n}{\varphi(\Delta_n) + \sqrt{\varphi(\Delta_n)^2 - 2Kg(\Delta_n)d_n}} \quad (n \geq 0) \tag{4.2}$$

(Moret [16], Yamamoto [32])

$$\leq \delta_n \equiv \frac{2g(t_n)d_n}{\varphi(t_n) + \sqrt{\varphi(t_n)^2 - 2Kg(t_n)d_n}} \quad (n \geq 0) \tag{4.3}$$

(Yamamoto [32])

$$= \frac{2(p_n/q_n)d_n}{1 + \sqrt{1 - 2KB_n d_n}} \quad (n \geq 0) \tag{4.4}$$

$$\leq \frac{2(p_n/q_n)d_n}{1 + \sqrt{1 - 2\sigma KB_n d_n}} \quad (n \geq 0) \tag{4.5}$$

$$\leq \frac{2(p_n/q_n)d_n}{1 + \sqrt{1 - 2h_n}} \quad (n \geq 0 \tag{4.6}$$

$$= \frac{t^* - t_n}{\nabla t_{n+1}} d_n \quad (n \geq 0) \tag{4.7}$$

$$\leq \frac{t^* - t_n}{\nabla t_n} d_{n-1} \quad (n \geq 1) \qquad \text{(Miel [15])} \tag{4.8}$$

$$= \frac{1}{\eta_{n-1}} \left(\frac{2(p_n/q_n)\eta_n}{1 + \sqrt{1-2h_n}}\right) d_{n-1} \quad (n \geq 1) \tag{4.9}$$

$$\leq t^* - t_n \quad (n \geq 0) \qquad \text{(Rheinboldt [25], Dennis [2])} \tag{4.10}$$

$$= \frac{2(p_n/q_n)\eta_n}{1 + \sqrt{1-2h_n}} \quad (n \geq 0) \quad . \tag{4.11}$$

(iv) Estimates

$$d_{n+1} \leq \frac{1}{g(\Delta_{n+1})} \left\{\frac{1}{2} Kd_n^2 + (m+M\Delta_n)d_n\right\} \leq \frac{\nabla t_{n+2}}{\nabla t_{n+1}} d_n \leq d_n$$

hold, where the last inequality may be replaced by the strict inequality $<$ if $d_n \neq 0$.

Proof. (i) An application of the majorant theory due to Rheinboldt [25] leads to the estimate

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n, \ n \geq 0 \quad ,$$

from which (i) follows by the standard argument and we obtain (4.10), since $\{t_n\}$ monotonically converges to t^*.

(ii) We have

$$F'(x_0) = A(x_0)\{I + A(x_0)^{-1}(F'(x_0) - A(x_0))\}^{-1}$$

and

$$\|A(x_o)^{-1}(F'(x_0) - A(x_0))\| \leq m < 1$$

by assumption so that

$$\|F'(x_0)^{-1}F(x_0)\| \leq \|F'(x_0)^{-1}A(x_0)\|\cdot\|A(x_0)^{-1}F(x_0)\| \leq \frac{\eta}{1-m} \quad ,$$

$$\|F'(x_0)^{-1}(F'(x) - F'(y))\| \leq \|F'(x_0)^{-1}A(x_0)\|\cdot\|A(x_0)^{-1}(F'(x) - F'(y))\|$$

$$\leq \frac{K}{1-m}\|x-y\| \quad , \quad x,y \in D_0 \quad ,$$

and

$$2\left(\frac{K}{1-m}\right)\left(\frac{\eta}{1-m}\right) \leq \frac{2\sigma K\eta}{(1 - \ell - m)^2} \leq 1 \quad .$$

Furthermore, let $\tilde{t}^*$ be the least solution of the equation $\tilde{f}(t) = \frac{1}{2}Kt^2 - (1-m)t + \eta = 0$. Then we have

$$\eta \leq \frac{\eta}{1-m} < \tilde{t}^* \leq t^*$$

since

$$f(t) \geq \tilde{f}(t) \ (t>0), \ \tilde{f}\left(\frac{\eta}{1-m}\right) = \frac{K}{2}\left(\frac{\eta}{1-m}\right)^2 > 0$$

and

$$\frac{\eta}{1-m} < \tilde{t}^{**} \quad \text{(the largest solution of } \tilde{f}(t) = 0) \quad .$$

Therefore we have $t^* - \eta \geq \tilde{t}^* - \frac{\eta}{1-m}$ and

$$\bar{S}(x_1, \tilde{t}^* - \frac{\eta}{1-m}) \subseteq \bar{S}(x_1, t^* - \eta) \subseteq D_0 \quad .$$

Consequently, the assumptions of Theorem 2.1 are satisfied by replacing K, η and t^* in

the theorem by $K/(1-m)$, $\eta/(1-m)$ and $\tilde{t}^*$, respectively. Hence we obtain from Theorem 2.1 (ii) that the solution is unique in the region $\tilde{S}$ defined in (4.1).

(iii) It is easy to see that $x^* \in \overline{S}_n \subseteq \overline{S}_{n-1}$. To obtain the bounds α_n, β_n, γ_n and δ_n, let

$$u_n(t) = \frac{1}{2} K_n t^2 + g(\Delta_n)^{-1}(m + M\Delta_n)t \quad ,$$

$$v_n(t) = \frac{1}{2} L_n t^2 + g(\Delta_n)^{-1}(m + M\Delta_n)t$$

$$w_n(t) = g(\Delta_n)^{-1} \{\frac{1}{2} Kt^2 + (m + M\Delta_n)t\} \quad ,$$

$$y_n(t) = g(t_n)^{-1} \{\frac{1}{2} Kt^2 + (m + Mt_n)t\} \quad ,$$

$$z_n(t) = g(t_n)^{-1} [\frac{1}{2} \sigma Kt^2 + \{m + (\sigma K-L)t_n\}t] + \nabla t_{n+1} - d_n \quad .$$

Then, as is easily seen, we have

$$\|x^* - x_{n+1}\| \leq u_n(\|x^* - x_n\|) \leq v_n(\|x^* - x_n\|) \leq w_n(\|x^* - x_n\|)$$

$$\leq y_n(\|x^* - x_n\|) \leq z_n(\|x^* - x_n\|) \quad .$$

Furthermore, observe that $t^* - t_n$ is the least solution of the equation $Z(t) \equiv z_n(t) - t + d_n = 0$. In fact, we have

$$(1 - \ell - m - \sigma Kt_n)^2 - 2\sigma Kg(t_n)\ \nabla t_{n+1}$$

$$= (1 - \ell - m - \sigma Kt_n)^2 - 2\sigma Kf(t_n)$$

$$= (1 - \ell - m)^2 - 2\sigma K\eta \geq 0 \quad ,$$

so that $Z(t) = 0$ has two positive solutions and

$$g(t_n)Z(t^* - t_n) = \frac{1}{2} \sigma K(t^* - t_n)^2 - (1 - \ell - m - \sigma Kt_n)(t^* - t_n) + f(t_n)$$

$$= 0 \quad . \tag{4.12}$$

This implies $Z(t^*-t_n) = 0$. Similarly we have $Z(t^{**} - t_n) = 0$, where t^{**} is the largest solution of $f(t) = 0$. Since we have already known by (i) that $\|x^* - x_n\| \leq t^* - t_n$, we can apply Theorem 3.1 to obtain

$$\|x^* - x_n\| \leq \alpha_n \leq \beta_n \leq \gamma_n \leq \delta_n \leq t^* - t_n \quad ,$$

where α_n, β_n, γ_n and δ_n denote the largest solutions of the equations

$$U_n(t) = u_n(t) - t + d_n = 0 \quad ,$$

$$V_n(t) = v_n(t) - t + d_n = 0 \quad ,$$

$$W_n(t) = w_n(t) - t + d_n = 0$$

and

$$Y_n(t) = y_n(t) - t + d_n = 0 \quad ,$$

respectively. Next, by induction on n, we shall prove that $t_{n+1} - t_n = \eta_n$. There is nothing to prove for $n = 0$. If $n \geq 1$ and $t_{k+1} - t_k = \eta_k$ is true for every $k \leq n-1$, then we have

$$\begin{aligned}
t_{n+1} - t_n &= g(t_n)^{-1} f(t_n) \\
&= g(t_n)^{-1}[\{f(t_n) - f(t_{n-1}) - f'(t_{n-1})\nabla t_n\} + \{f'(t_{n-1}) + g(t_{n-1})\}\nabla t_n] \\
&= g(t_n)^{-1}[\tfrac{1}{2}\sigma K(\nabla t_n)^2 + \{m + (\sigma K - L)t_{n-1}\}\nabla t_n] \\
&= g\Big(\sum_{j=0}^{n-1} \eta_j\Big)^{-1}\Big[\tfrac{1}{2}\sigma K\eta_{n-1}^2 + \Big\{m + (\sigma K - L)\sum_{j=0}^{n-2} \eta_j\Big\}\eta_{n-1}\Big] \\
&= p_n^{-1}\{\tfrac{1}{2}\sigma K\eta_{n-1}^2 + (p_{n-1} - q_{n-1})\eta_{n-1}\} \\
&= \eta_n \quad ,
\end{aligned}$$

where we understand that $\sum_{j=0}^{-1} \eta_j = 0$. Furthermore, we have

$$\begin{aligned}
h_n &= \sigma K(p_n/q_n^2)\eta_n \\
&= \sigma K[\tfrac{1}{2}\sigma K(\nabla t_n)^2 + \{m + (\sigma K - L)t_{n-1}\}\nabla t_n]/(1 - \ell - m - \sigma K t_n)^2 \\
&= \sigma K[\tfrac{1}{2}\sigma K(\nabla t_n)^2 + \{m + (\sigma K - L)t_{n-1}\}\nabla t_n]/(1 - \ell - m - \sigma K t_n)^2 \\
&= \sigma K[f(t_n) - f(t_{n-1}) - f'(t_{n-1})\nabla t_n + \{f'(t_{n-1}) + g(t_{n-1})\}\nabla t_n]/(1 - \ell - m - \sigma K t_n)^2
\end{aligned}$$

$$= \sigma K f(t_n)/(1 - \ell - m - \sigma K t_n)^2$$

$$= \frac{1}{2}\{(\sigma K t_n)^2 - 2\sigma K(1 - \ell - m)t_n + 2\sigma K\eta\}/(1 - \ell - m - \sigma K t_n)^2$$

$$= \frac{1}{2}\{(1 - \ell - m - \sigma K t_n)^2 + 2\sigma K\eta - (1 - \ell - m)^2\}/(1 - \ell - m - \sigma K t_n)^2$$

$$\leq \frac{1}{2}$$

and, by (4.12) and $z(t^{**} - t_n) = 0$,

$$t^* - t_n = \frac{1 - \ell - m - \sigma K t_n - \sqrt{(1-\ell-m-\sigma K t_n)^2 - 2\sigma K f(t_n)}}{\sigma K}$$

$$= \frac{q_n - \sqrt{q_n^2 - 2\sigma K g(t_n)\nabla t_{n+1}}}{\sigma K}$$

$$= \frac{q_n - \sqrt{q_n^2 - 2\sigma K p_n \eta_n}}{\sigma K}$$

$$= \frac{2(p_n/q_n)\eta_n}{1 + \sqrt{1-2h_n}} .$$

This leads to the estimates

$$\delta_n = \frac{2(p_n/q_n)d_n}{1 + \sqrt{1-2K(p_n/q_n^2)d_n}} \quad (n \geq 0)$$

$$= \frac{2(p_n/q_n)d_n}{1 + \sqrt{1-2KB_n d_n}} \quad (n \geq 0)$$

$$\leq \frac{2(p_n/q_n)d_n}{1 + \sqrt{1 - 2\sigma KB_n d_n}} \quad (n \geq 0)$$

$$\leq \frac{2(p_n/q_n)d_n}{1 + \sqrt{1-2h_n}} \quad (n \geq 0)$$

$$= \frac{2(p_n/q_n)\eta_n}{1 + \sqrt{1-2h_n}} \frac{d_n}{\eta_n} \quad (n \geq 0)$$

$$= (t^* - t_n) \frac{d_n}{\nabla t_{n+1}} \quad (n \geq 0)$$

$$\leq (t^* - t_n) \cdot \frac{d_{n-1}}{\nabla t_n} \ (n \geq 1)$$

$$= \frac{2(p_n/q_n)\eta_n}{1 + \sqrt{1-2h_n}} \cdot \frac{d_{n-1}}{\eta_{n-1}} \ (n \geq 1)$$

$$\leq \frac{2(p_n/q_n)\eta_n}{1 + \sqrt{1-2h_n}} \ (n \geq 0)$$

$$= t^* - t_n \ (n \geq 0)$$

$$\leq 2(p_n/q_n)\eta_n \ (n \geq 0)$$

$$= q_n^{-1}\{\sigma K\eta_{n-1}^2 + 2(p_{n-1} - q_{n-1})\eta_{n-1}\} \ (n \geq 0) \quad ,$$

where we have used Miel's result [14] $d_n/\nabla t_{n+1} \leq d_{n-1}/\nabla t_n$ $(n \geq 1)$.

(iv) The statement (iv) is proved in [32].

Q.E.D.

It is clear that Theorem 4.1 generalizes Theorem 2.1, although the latter was used in the proof of the former. As another result obtained from Theorem 4.1, we have the following corollary, too.

Corollary 4.1.1. Consider the modified Newton method

$$x_{n+1} = x_n - F'(x_0)^{-1}F(x_n) \quad , \quad n \geq 0 \quad , \tag{4.13}$$

where we assume the following:

$$x_0 \in D_0,\ F'(x_0)^{-1} \text{ exists ,}$$

$$\|F'(x_0)^{-1}(F'(x) - F'(y))\| \leq K\|x-y\|,\ x,y \in D_0 \ ,$$

$$\eta = \|F'(x_0)^{-1}F(x_0)\| > 0,\ h = K\eta \leq \frac{1}{2} \ ,$$

$$t^* = (1 - \sqrt{1-2h})/K,\ t^{**} = (1 + \sqrt{1-2h})/K \ ,$$

$$\bar{S} = \bar{S}(x_1,\ t^* - \eta) \subseteq D_0 \ .$$

Then:

(i) The iteration (4.13) is well defined fo every $n \geq 0$, $x_n \in S$ for $n \geq 1$ and $\{x_n\}$ converges to a solution of (1.1).

(ii) The solution is unique in

$$\tilde{S} = \begin{cases} S(x_0, t^{**}) \cap D_0 & (2h < 1) \\ \bar{S}(x_0, t^{**}) \cap D_0 & (2h = 1) \end{cases} .$$

(iii) Define the sequence $\{t_n\}$ by

$$t_0 = 0, \; t_{n+1} = \frac{1}{2} Kt_n^2 + n, \; n \geq 0 .$$

Put $\bar{S}_0 = \bar{S}$, $\bar{S}_n = \bar{S}(x_n, t^* - t_n)$ $(n \geq 1)$,

$$K_n = \sup_{\substack{x,y \in \bar{S}_n \\ x \neq y}} \frac{\|F'(x_0)^{-1}(F'(x) - F'(y))\|}{\|x - y\|} \; (n \geq 0) .$$

Then we have

$$\begin{aligned}
\|x^* - x_n\| &\leq \frac{2d_n}{1 - K\Delta_n + \sqrt{(1-K\Delta_n)^2 - 2K_n d_n}} \; (n \geq 0) \\
&\leq \frac{2d_n}{1 - K\Delta_n + \sqrt{(1-K\Delta_n)^2 - 2K_0 d_n}} \; (n \geq 0) \\
&\leq \frac{2d_n}{1 - K\Delta_n + \sqrt{(1-K\Delta_n)^2 - 2Kd_n}} \\
&\leq \frac{2d_n}{1 - Kt_n + \sqrt{(1-Kt_n)^2 - 2Kd_n}} \; (n \geq 0) \\
&\leq \frac{t^* - t_n}{\nabla t_{n+1}} \; (n \geq 0) \\
&\leq \frac{t^* - t_n}{\nabla t_n} d_{n-1} \; (n \geq 1) \\
&\leq t^* - t_n \; (n \geq 0) \\
&= \frac{2(1-Kt_n)^{-1}\nabla t_{n+1}}{1 + \sqrt{1-2K(1-Kt_n)^{-2}\nabla t_{n+1}}} \; (n \geq 0) .
\end{aligned}$$

Finally we remark that the approach employed in this paper is also applicable to other types of iterations.

REFERENCES

[1] J. E. Dennis, On the Kantorovich hypothesis for Newton's method. SIAM J. Numer. Anal., 6(1969), 493-507.

[2] J. E. Dennis, On the convergence of Newton-like methods. Numerical Methods for Nonlinear Algebraic Equations (ed. P. Rabinowitz), Gordon and Breach Science Publishers, New York, 1970, 163-181.

[3] J. E. Dennis, Toward a unified convergence theory for Newton-like methods. Nonlinear Functional Analysis and Applications (ed. L. B. Rall), Academic Press, New York, 1971, 425-472.

[4] P. Deuflhard and G. Heindl, Affine invariant convergence theorems for Newton's method and extensions to related methods. SIAM J. Numer. Anal., 16(1979), 1-10.

[5] B. Döring, Über das Newtonsche Näherungsverfahren. Math. Phys. Sem.-Ber., 16(1969), 27-40.

[6] W. B. Gragg and R. A. Tapia, Optimal error bounds for the Newton-Kantorovich theorem. SIAM J. Numer. Anal., 11(1974), 10-13.

[7] L. V. Kantorovich, On Newton's method for functional equations. Dokl. Akad. Nauk. SSSR, 59(1948), 1237-1240.

[8] L. V. Kantorovich, The majorant principle and Newton's method. Dokl. Akad. Nauk. SSSR, 76(1951), 17-20.

[9] L. V. Kantorovich and G. P. Akilov, Functional Analysis in Normed Spaces. Pergamon Press, Oxford, 1964.

[10] H. J. Kornstaedt, Functional ungleichungen und Iterationsverfahren. Aequat. Math., 13(1975), 21-45.

[11] P. Lancaster, Error analysis for the Newton-Raphson method. Numer. Math., 9(1966), 55-68.

[12] G. J. Miel, The Kantorovich theorem with optimal error bounds. Amer. Math. Mon., 86(1979), 212-215.

[13] G. J. Miel, Unified error analysis for Newton-type methods. Numer. Math., 33(1979), 391-396.

[14] G. J. Miel, Majorizing sequences and error bounds for iterative methods. Math. Comp., 34(1980), 185-202.

[15] G. J. Miel, An updated version of the Kantorovich theorem for Newton's method. Computing, 27(1981), 237-244.

[16] I. Moret, A note on Newton type iterative methods. Computing, 33(1984), 65-73.

[17] J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York, 1970.

[18] A. M. Ostrowski, La méthod de Newton dans les espaces de Banach. C. R. Acad. Sci. Paris, 27(A)(1971), 1251-1253.

[19] A. M. Ostrowski, Solution of Equations in Euclidean and Banach Spaces. Academic Press, New York, 1973.

[20] F. A. Potra, On the convergence of a class of Newton-like methods. Iterative Solution of Nonlinear Systems of Equations (eds. R. Ansarge and W. Törnig), Lecture Notes in Math. 953, Springer Verlag, Berlin-Heidelberg-New York, 1982.

[21] F. A. Potra, On the aposteriori error estimates for Newton's method. Beiträge zur Numer. Math., 12(1984), 125-138.

[22] F. A. Potra and V. Pták, Sharp error bounds for Newton's process. Numer. Math., 34(1980), 63-72.

[23] L. B. Rall, Computational Solution of Nonlinear Operator Equations. Krieger, Huntington, NY, 1979.

[24] L. B. Rall and R. A. Tapia, The Kantorovich theorem and error estimates for Newton's method. MRC Technical Summary Report, No. 1043, University of Wisconsin-Madison, 1970.

[25] W. C. Rheinboldt, A unified convergence theory for a class of iterative process. SIAM J. Numer. Anal., 5(1968), 42-63.

[26] J. W. Schmidt, Regular-falsi-Verfahren mit konsistenter Steigung und Majorantenprinzip. Periodica Math. Hungarica, 5(1974), 187-193.

[27] J. W. Schmidt, Untere Fehrerschranken für Regular-falsi-Verfahren. Periodica Math. Hungarica, 9(1978), 241-247.

[28] R. A. Tapia, The Kantorovich theorem for Newton's method. Amer. Math. Mon., 78(1971), 389-392.

[29] T. Yamamoto, Error bounds for Newton's process derived from the Kantorovich theorem. Japan J. Appl. Math., 2(1985), to appear.

[30] T. Yamamoto, Error bounds for Newton's iterates derived from the Kantorovich theorem. Numer. Math., in press.

[31] T. Yamamoto, A unified derivation of several error bounds for Newton's process. J. Comp. Appl. Math., 12 & 13(1985), 179-191.

[32] T. Yamamoto, Error bounds for Newton-like methods under Kantorovich type assumptions. MRC Technical Summary Report, No. 2846, University of Wisconsin-Madison, 1985.

[33] T. Yamamoto, A convergence theorem for Newton's method in Banach spaces. MRC Technical Summary Report, No. 2879, University of Wisconsin-Madison, 1985.

[34] T. Yamamoto, Error bounds for Newton's method under the Kantorovich assumptions. Proceedings of the Conference "New Directions in Applied and Computational Math." held in the University of Wyoming, 1985, to appear.

[35] T. Yamamoto, A method for finding sharp error bounds for Newton's method under the Kantorovich assumptions. preprint.

TY/jvs

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER #2904	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A CONVERGENCE THEOREM FOR NEWTON-LIKE METHODS IN BANACH SPACES		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Tetsuro Yamamoto		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53705		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 3 - Numerical Analysis and Scientific Computing
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE January 1986 13. NUMBER OF PAGES 16
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

A convergence theorem, error estimates, Newton-like methods, Kantorovich type assumptions

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A convergence theorem for Newton-like methods in Banach spaces is given, which improves results of Rheinboldt [25], Dennis [2], Miel [13, 14] and Moret [16] and includes as a special case an updated version of the Kantorovich theorem for the Newton method given in previous papers [33-35]. Error bounds obtained in [32] are also improved. This paper unifies the study of finding sharp error bounds for Newton-like methods under Kantorovich type assumptions.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

END
FILMED

6-86

DTIC